Evaluating Correlation Between Measurement Samples in Reverberation Chambers Using Clustering

Carnot L. Nogueira [#$1], Kate A. Remley [#2], Robert D. Horansky [#3]
[#]National Institute of Standards and Technology (NIST), Boulder, CO 80305 USA
[$]Department of Physics, University of Colorado at Boulder, Boulder, CO 80309 USA
[1]carnot.nogueira@nist.gov, [2]kate.remley@nist.gov, [3]robert.horansky@nist.gov

***Abstract* — Traditionally, in reverberation chambers (RC) measurement autocorrelation or correlation-matrix methods have been applied to evaluate measurement correlation. In this article, we introduce the use of clustering based on correlative distance to group correlated measurements. We apply the method to measurements taken in an RC using one and two paddles to stir the electromagnetic fields and applying decreasing angular steps between consecutive paddles' positions. The results – using varying correlation threshold values – demonstrate that the method calculates the number of effective samples and allows discerning outliers, i.e., uncorrelated measurements, and clusters of correlated measurements. This calculation method, if verified, will allow non-sequential stir sequence design and, thereby, reduce testing time.**

***Keywords* — Correlation, Pearson correlation coefficient (PCC), reverberation chambers (RC), mode-stirring samples, correlative distance, clustering analysis, adjacency matrix.**

I. Introduction

Since mobile phones became popular, in the 1990s, the use of wireless communication devices has grown at a very fast rate and today pervades many aspects of modern life. Wireless computing, uncrewed vehicles, internet-of-things (IoT) and industrial IoT (IIoT), handheld mobile devices, and radar communications are only a few examples of wireless communication applications [1].

Reverberation chambers (RCs) can be used to test both small and large wireless devices. RCs are typically rectangular parallelepipedic shielded resonating cavities that support a large set of resonant electromagnetic (EM) waves. RCs are usually equipped with paddles that stir the EM fields preventing the formation of static minima and maxima regions of field amplitudes [2]-[4]. Other mechanisms can be used to stir the fields and to move devices under test (DUT) to different positions (e.g., turntables, positioners). Other procedures, such as changes in antenna orientation or switching between antennas, can also be used to change or sample the EM fields at different physical locations. EM field boundary conditions are also significantly changed by the presence of RF absorbers.

Correlation between mode-stirring samples has the adverse effect of reducing measurement efficiency but can be a necessary trade-off for the accurate measurement of modulated signals [4]. When correlation exists between RC measurement samples the experiment must be designed to avoid an insufficient number of uncorrelated samples necessary for a desired measurement uncertainty [5]. Consequently, to conduct accurate and efficient RC experiments, correlations between measurement samples must be evaluated.

Evaluation of correlation in RC-related experiments has been a concern for wireless industry and researchers for more than 30 years [6] and the efficiency of stirrers has been assessed using different approaches. IEC 61000-4-21 adopts a method that uses a power-based thresholded correlation coefficient and assumes that the stirrer positions are distributed evenly [7], [8]. Pfennig and Krauthäuser developed a general method based on the PCC for each frequency and the stirrer positions. The positions are regarded as independent if the PCC is less than a correlation threshold, algorithms then identify groups of independent stirrer positions [8]-[10]. Gradoni and co-workers evaluated mode-stirring efficiency in RCs using uniformly distributed grid of points inside the chamber. Their results showed that circular correlation methods overestimate stirrer efficiency [11]. A correlation matrix-based method using complex S_{21}-parameters was introduced by Pirkl et al. [5], in the method independent observations obtained from a range of frequencies were used to form Pearson correlation matrices.

The method introduced in this paper uses the concept of *correlative distance* between S_{21}-parameter measurements in the set of real numbers ($\mathbb{R}$) to build an adjacency matrix and group correlated data into clusters. It allows the calculation of the number of effective samples for any given threshold (r_{lim}) applied to the PCC through the identification of both outliers, i.e., measurements that have no correlation to any other measurement, and clusters, i.e., groups of two or more correlated measurements.

II. Derivation of the Method

Any two sets of n real variables $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ and $\boldsymbol{Y} = \{Y_1, Y_2, \ldots, Y_n\}$ can be standardized into variables $\boldsymbol{x} = \{x_i\}$ and $\boldsymbol{y} = \{y_i\}$, with means $\bar{x} = 0$ and $\bar{y} = 0$, and with standard deviations computed with the n-method given by $s_x = 1.0$ and $s_y = 1.0$ [12]. The PCC ($r_{X,Y}$) between $\boldsymbol{X}$ and $\boldsymbol{Y}$ data sets can then be calculated from:

$$r_{X,Y} = \frac{1}{n}\sum x_i y_i \tag{1}$$

Using the equation above, the squared Euclidean distance d between x and y can be proven to be [12, 13]:

$$d^2 = \sum(x_i - y_i)^2 = 2n(1 - r_{X,Y}) \tag{2}$$

The squared distance d, as shown above, is proportional to $1 - r_{X,Y}$ and the square root of the last term in (2) is the *correlative distance* between x and y [12].

Correlation between measurements in RCs can be calculated using complex S_{21}-parameters recorded over pre-defined frequency intervals. PCC between position 1 and position 2 of

any two different configurations (throughout the paper the term "configuration" refers to a specific position of the paddles that can be used to generate one sample) can be calculated as [4]:

$$\rho(f) = \left| \frac{\Sigma\left[(S_{21,1} - \overline{S_{21,1}})(S_{21,2} - \overline{S_{21,2}})^*\right]}{\sqrt{\Sigma\left[\left|(S_{21,1} - \overline{S_{21,1}})\right|^2\right]}\sqrt{\Sigma\left[\left|(S_{21,2} - \overline{S_{21,2}})\right|^2\right]}} \right| \quad (3)$$

where (*) is the complex conjugate. An *n*-tuple sequence of sampled frequency-dependent complex S_{21}-parameters $\{S_{21,1} = a_1 + ib_1, S_{21,2} = a_2 + ib_2, \ldots, S_{21,n} = a_n + ib_n\} \in \mathbb{C}^n$ can be rewritten as a $2n$-tuple sequence $\{a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n\} \in \mathbb{R}^{2n}$. In $\mathbb{R}^{2n}$, PCC between configurations can then be computed using (1) and although the numeric values of the PCC values calculated in $\mathbb{C}^n$ using (3) are not the same as the ones calculated in $\mathbb{R}^{2n}$ using (1), the values are very close.

The application of (2) to complex S_{21}-parameter stirring-sequence samples represented in $\mathbb{R}^{2n}$ allows the calculation of the *correlative distance matrix* (CDM), where each element corresponds to the numerical value of the correlative distance between measurements *i* and *j*:

$$d_{i,j} = \sqrt{2n(1 - r_{i,j})} \quad (4)$$

CDMs can be used to determine whether the correlation between any pair of measurements is beyond some *correlative distance threshold* (d_{thrs}) that can also be computed using (4) and the predefined PCC threshold (r_{thrs}). Instead of using CDMs with the correlative distance values, binary CDMs (with entries equal to 1 when correlation between measurements is above the threshold and 0 if it is not) can be adopted to form what we term *correlative distance adjacency matrices*. The analysis of these adjacency matrices, for a given threshold, allows the identification of measured samples that are not correlated to any other (outliers) and of groups of two or more samples that are correlated (clusters). The number of effective samples (N_{eff}) can then be determined by adding the number of outliers (N_{outls}) and the number of clusters (N_{clust}):

$$N_{\text{eff}} = N_{\text{outls}} + N_{\text{clust}} \quad (5)$$

It is assumed that all measurements within a cluster, since they are correlated to each other, correspond to one effective sample only.

III. Results from an RC with Two Paddles

Results presented in this section are based on measurements performed in a large, double-paddled reverberation chamber equipped with a turntable and a vertical positioner and loaded with four absorbers. The chamber (dimensions 4.27 m x 3.65 m x 2.90 m) has a vertical and a horizontal cylindrical paddle (lengths 2.50 m and 3.30 m, respectively, and both with 0.55 m radius to the tip of paddle). A dual-ridged horn antenna (bandwidth 1 GHz – 18 GHz) was used to transmit the signals and a discone antenna to receive (bandwidth 0.65 GHz - 3.5 GHz), Fig. 1 shows both antennas inside the chamber. In the configuration of the chamber, a total of 360 paddle positions (1° steps) were used in the vertical paddle rotation and 2 positions only in the horizontal paddle (0° and 180°), whereas the turntable and the positioner were kept at fixed positions.

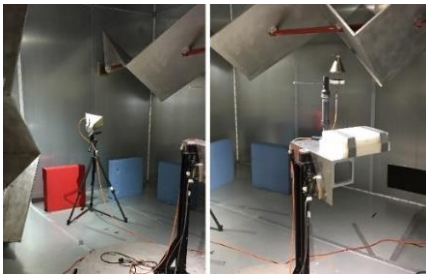

Figure 1: Pictures of the transmit antenna (horn antenna, left) and receive antenna (discone, right) inside the chamber.

The following parameters were used in the vector network analyzer (VNA): center frequency 2.575 GHz, bandwidth 3.85 GHz, and 1 kHz IF bandwidth; frequencies, therefore, ranged from 650 MHz to 4.5 GHz, with frequency points read at 0.25 MHz intervals for a total of 15401 points. Coherence bandwidth (CBW) for the full frequency range is 3.628 MHz and for center frequencies 1.5, 2.5, 3.5 GHz and bandwidth 1.0 GHz, CBWs are 3.603, 3.517, 3.505 MHz, respectively.

For this set-up, using the entire frequency range and the vertical paddle only, the coherence angles computed using the autocovariance method described in [5] for r_{thrs} = 0.37, 0.5, and 0.707 are, respectively, approximately 21°, 18°, and 11° corresponding to N_{eff} = 17 (360°/21° = 17), 20 (360°/18° = 20), and 33 (360°/11° = 33). For the same thresholds, the clustering approach introduced here yields N_{eff} = 1, 20, and 50, respectively. The discrepancy between the results is under investigation. It is worth emphasizing that the autocorrelation method is restricted to physically adjacent paddle angles, while no such restrictions exist in the clustering approach, besides, correlation methods have been reported to overestimate stirrer efficiency [11].

In the examples below, the clustering method estimates the average N_{eff} over the entire range of frequencies, nonetheless, any subset of frequencies with any number of frequency points can be used. The complex S_{21}-parameters were stored in a *c*-configurations' by *f*-frequencies' matrix $\mathbf{X} \in \mathbb{C}^{\text{c x f}}$.

Example 1 In this example, the vertical paddle was at a fixed position, while in the horizontal paddle only angles spaced 20° were measured, resulting in a total of 18 configurations (c = 18), corresponding to angles 0°, 20°, 40°, … 340°. Using all frequency points, a complex matrix $\mathbf{X} \in \mathbb{C}^{18 \text{ x } 15401}$ was used in the analysis. For a threshold distance r_{thrs} = 0.470, and real matrix $\mathbf{X} \in \mathbb{C}^{18 \text{ x } (2 \cdot 15401)}$, the correlative distance threshold is d_{thrs} = 180.7 (4) and the resulting correlative distance adjacency matrix is shown in Fig. 2.

As expected, since $r_{X,Y} = r_{Y,X}$, the matrix is symmetric. The diagonal is not represented since any configuration is correlated to itself (PCC = 1.0). Outliers are represented by rows (or columns) with no markers, i.e., no correlation with any other configuration. In Fig. 2, configurations 3 and 4 are outliers. Clusters are pairs or larger groups of configurations where any element is correlated to at least one other element in the cluster. In Fig. 2, markers on row 1 (or column 1) show correlation

between configurations 1 and 2 and also 1 and 18; group {1, 2, 18}, therefore, together form a cluster. Another cluster, with 2 configurations only, is formed by configurations 5 and 6. A third cluster is defined by configurations represented in rows 7 to 10: {7, 8, 9, 10}. Cluster 4 is composed of 11, 12, 13, and 14. Finally, the last cluster includes configurations 15, 16, 17. The number of effective samples is N_{eff} = 7 (5), where two are outliers (N_{outls} = 2), and five are clusters (N_{clust} = 5). Since all configurations in each cluster can be replaced by one of the configurations, clusters can be used to design stirring sequences with non-equally spaced samples. It is also important to emphasize that the analysis, although presented here for the whole frequency range, can be used to any arbitrary subset.

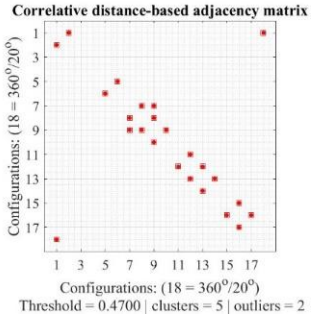

Figure 2: Correlative distance adjacency matrix: the red markers represent configurations with correlation above r_{thrs} between measurements. In row 1, the red markers indicate correlation between configuration 1 and 2 and between 1 and 18; together, configurations 1, 2, and 3 form a cluster.

Example 2 The horizontal paddle angles are now spaced 6°, resulting in 60 configurations, whereas the vertical paddle remains at a fixed position. Matrix **X** is now 60 by 15401 and for the threshold distance r_{thrs} = 0.729 the following adjacency matrix results (Fig. 3), with 12 clusters and 25 outliers, for a total number of effective samples N_{eff} = 37.

Correlation between successive positions, e.g., {18, 19}, {34, 35}, and {59, 60}, with the markers close to the main diagonal of the correlative distance matrix, is due to the fact that the paddle moves to successive neighboring positions, as the angle increases, such as {18, 19} = {108°, 114°}, and neighboring positions tend to be correlated, because the distance between them is the 6° step used in the analysis. Were other stirring mechanisms being used, off-diagonal correlations could occur, creating clustered configurations in any region of the adjacency matrix.

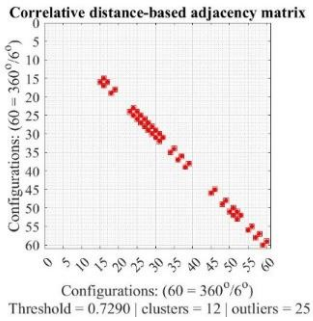

Figure 3: Correlative distance adjacency matrix: angle step size 6°, resulting in 60 configurations with paddle angles 0°, 6°, 12°,… 354°.

A study showing the variation of the number of effective samples as the threshold r_{thrs} varies from 0.0 to 1.0 shows interesting aspects of the chamber and stirring mechanisms behavior (Fig. 4).

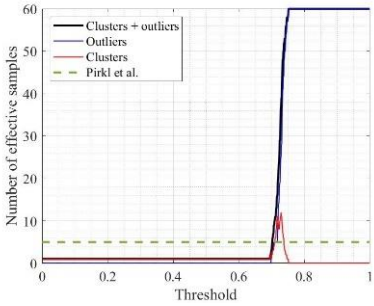

Figure 4: Variation of the number of clusters and effective samples as the threshold distance varies from 0.0 to 1.0 for a stirrer step size of 6°.

As shown on Fig. 4, for low values of the threshold, less than approximately 0.7, there are no outliers (blue line) and all measurements are correlated, forming a big cluster containing all measurements (black and red lines). As the threshold increases, in the [0.70, 0.729] interval, the number of clusters and outliers then increase up to the point where the number of clusters reaches a maximum, at r_{thrs} = 0.729 (Fig. 3). Beyond that point, i.e., the point where the number of clusters reaches a maximum, the increase in r_{thrs} begin to *extract* mode-stirring samples from the clusters, as the correlation condition gets more restrictive. The number of clusters then decreases to zero, as the number of outliers equals the total number of mode-stirring samples in the analysis (60, corresponding to 6° steps). This result is expected, since beyond some threshold values (in Fig. 4. r_{thrs} = 0.750) all samples will be uncorrelated. The application of the proposed method with r_{thrs} varying from 0.0 to 1.0 allows the understanding of how the mode-stirring samples behave in terms of being correlated or not correlated. The analysis of the clusters allows the evaluation of the efficiency of the stirring mechanisms.

The method proposed here is compared to the approach used by Pirkl et al. [5]. The method estimates the number of effective samples from the quotient of the squared number of configurations (N) and the sum of the complex correlation matrix entries ($r_{i,j}$) absolute values squared (Eq. 12 in [5]):

$$N_{eff} = \frac{N^2}{\sum_{i,j=1}^{N} |r_{ij}|^2} \quad (6)$$

As depicted in Fig. 4, while Pirkl's method shows no variation as the threshold increases – and this happens because the threshold value is not a parameter in (6) – the method presented here shows no outliers and only one cluster (containing all samples) before, approximately, r_{thrs} = 0.7. Beyond this point, the method captures a steep increase in the number of effective samples as r_{thrs} varies from approximately 0.700 to 0.75; the method then reaches and stays at the maximum number of effective samples (N_{eff} = 60) in the interval $0.75 < r_{thrs} \leq 1.0$. The sharp increase in the number of effective samples and the permanence in the highest N_{eff} plateau is a consequence of the

stricter requirement – very high r_{thrs} (3) – for any two samples to be considered correlated; in the limit, for very small angular difference between measurements, only at r_{thrs} = 1.0 the measurements will be rendered uncorrelated. In the interval 0.7 < r_{thrs} < 0.75, the method shows an increase and then a decrease in the number of clusters. The maximum number of clusters (red line) is 12 and occurs at r_{thrs} = 0.729 (correlative distance adjacency matrix shown in Fig. 3).

Example 3 In this example, both paddles are used: the vertical paddle with the most refined angular step (1º, 360 positions) and the horizontal paddle at 0° and 180° (2 positions), resulting on a total of 720 mode-stirring samples. The variation of the number of effective samples as the threshold increases is shown in Fig. 5.

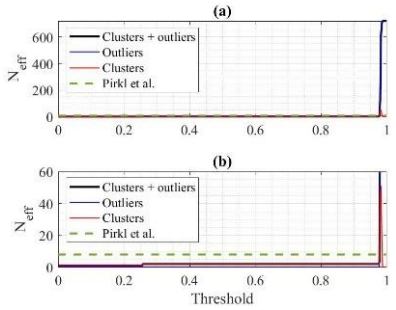

Figure 5: (a) Variation of the number of clusters and effective samples using both the vertical (360 positions, 1º steps) and the horizontal paddle (2 positions, 0º and 180º). (b) Detail of the y-interval [0, 60] ($0 \leq N_{eff} \leq 60$) showing the *decoupling* of the horizontal paddle positions at r_{thrs} = 0.255 and the increase and decrease in the number of clusters at r_{thrs} > 0.97.

As expected, the number of effective samples shows a steep increase as the threshold approaches 1.0 (r_{thrs} > 0.97, approximately) to reach the maximum number of effective samples N_{eff} = 720. The formation of clusters and the steep increase in the number of effective samples close to r_{thrs} = 1.0 indicates that the mode-stirring samples are very strongly correlated, which is a consequence of the low step angle used in the analysis (only 1º). Since the difference between consecutive positions is small, only when the correlation threshold is very close to 1.0 the measurements are perceived as uncorrelated. In Fig. 5b (detail of the vertical axis [0, 60] interval) the increase from 1 to 2 effective samples at r_{thrs} = 0.255 shows a *decoupling* of the 360 vertical paddle positions associated with position 0º of the horizontal paddle position and the 360 vertical paddle positions associated with position 180º in the horizontal paddle. This demonstrates that the 180º difference in the horizontal paddle positions creates uncorrelated sets of configurations (in the vertical paddle positions) detectable at a very low threshold: r_{thrs} = 0.255. The increase and decrease in the number of clusters is also readily visible at 0.97 < r_{thrs} < 0.99. Pirkl's approach remains constant, as the method (5) does not account for variations in r_{thrs}.

IV. Conclusion

We presented a method that allows a thorough analysis of correlation between mode-stirring samples generated in RC environments equipped with paddles. The method is based on the correlative distance between S_{21}-parameter n-tuples measured from mode-stirring samples (in complex space $\mathbb{C}$) represented as 2n-tuples of real numbers (in real space $\mathbb{R}$). The distance between measurements represented in $\mathbb{R}$ can be directly associated with PCC and a clustering scheme can be employed to analyze correlation.

Examples to validate the proposed method showed that it can be used to calculate the number of total effective samples and demonstrated that the clusters formed in the analysis group correlated mode-stirring samples' measurements. Clusters can then be used to implement stirring sequences with non-equally spaced paddle angles.

References

[1] "Global Digital Overview". Datareportal – Global Digital Insights. Retrieved March 23, 2020

[2] C. L. Holloway, D. A. Hill, J. M. Ladbury, P. F. Wilson, G. Koepke and J. Coder, "On the Use of Reverberation Chambers to Simulate a Rician Radio Environment for the Testing of Wireless Devices," in IEEE Transactions on Antennas and Propagation, vol. 54, no. 11, pp. 3167-3177, Nov. 2006, doi: 10.1109/TAP.2006.883987.

[3] R. D. Horansky and K. A. Remley, (2019), "Flexibility in over-the-air testing of receiver sensitivity with reverberation chambers". IET Microw. Antennas Propag., 13: 2590-2597

[4] K. A. Remley, C. Wang and R. D. Horansky, (2021), "Over-the-air testing of wireless devices in heavily loaded reverberation chambers," The Institution of Engineering and Technology, London, [online], https:// tsapps. nist. gov/ publication/ get_pdf. cfm?pub_id=928590 (Accessed December 29, 2021)

[5] R. J. Pirkl, K. A. Remley and C. S. Lötbäck Patane, "Reverberation Chamber Measurement Correlation," in IEEE Transactions on Electromagnetic Compatibility, vol. 54, no. 3, pp. 533-545, June 2012.

[6] B. Boverie, "Mode-Stirred Chamber Field Statistics: Correlation Widths," Oral presentation, Reverberating Chamber User's Group Meeting, Boulder, CO, Aug. 1991.

[7] "Electromagnetic Compatibility (EMC)—Part 4.21: Testing and Measurement Techniques—Reverberation Chamber Test Methods," International Electrotechnical Commission, IEC 61000-4-21, 2011.

[8] S. Pfennig and H. G. Krauthäuser, "A general method for determining the number of independent stirrer positions in reverberation chambers," in Proc. Int. Symp. Electromagn. Compat., Sep. 2012, pp. 1–6.

[9] S. Pfennig and H. G. Krauthäuser, "Comparison of methods for determining the number of independent stirrer positions in reverberation chambers," in Proc. Int. Symp. Electromagn. Compat., Sep. 2013, pp. 431–436.

[10] S. Pfennig, "A General Method for Determining the Independent Stirrer Positions in Reverberation Chambers: Adjusting the Correlation Threshold," in IEEE Transactions on Electromagnetic Compatibility, vol. 58, no. 4, pp. 1252-1258, Aug. 2016.

[11] G. Gradoni, V. Mariani Primiani, and F. Moglie, "Reverberation chamber as a multivariate process: FDTD evaluation of correlation matrix and independent positions," Progress In Electromagnetics Research, vol. 133, pp. 217–234, 2013.

[12] P. Vos, "Pearson's correlation between three variables; Using students' basic knowledge of geometry for an exercise in mathematical statistics," International Journal of Mathematical Education in Science and Technology. 40. 533-541. 10.1080/00207390802419578.

[13] I. Borg and P. J. F. Groenen, "How to Obtain Proximities, in Modern Multidimensional Scaling: Theory and Applications," Springer New York, New York, NY, 978-0-387-28981-6, 2005.